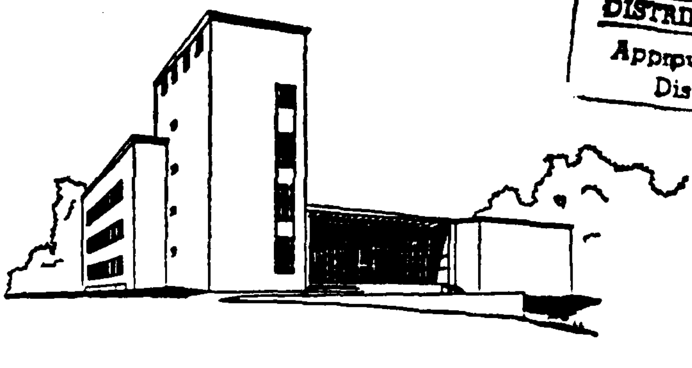
1

Carnegie Mellon University

PITTSBURGH, PENNSYLVANIA 15213

GRADUATE SCHOOL OF INDUSTRIAL ADMINISTRATION

WILLIAM LARIMER MELLON, FOUNDER



1

W.P. #87-88-40

Management Science Research Report No. MSRR 545

ON THE CONVEX HULL
OF THE UNION OF CERTAIN POLYHEDRA

by

Egon Balas

Carnegie Mellon University

May 1988

Revised July 1988

The research underlying this report was supported by Grant ECS-8601660 of the National Science Foundation, Contract N00014-85-K-0198 with the Office of Naval Research, and Grant AFOSR-870292 of the Air Force Office of Scientific Research. Reproduction in whole or in part is permitted for any purpose of the U.S. Government.

Management Science Research Group
Graduate School of Industrial Administration
Carnegie Mellon University
Pittsburgh, PA 15213

Abstract

This document ~~We~~ considers a finite collection of polyhedra whose defining linear systems differ only in their right hand sides. Jeroslow [5] and Blair [4] specified conditions under which the convex hull of the union of these polyhedra is defined by a system whose left hand side is the common left hand side of the individual systems, and whose right hand side is a convex combination of the individual right hand sides. ~~We~~ give a new sufficient condition for this property to hold, which is often easier to recognize. In particular, it is ~~we~~ shown that the condition is satisfied for polyhedra whose defining systems involve the node-arc incidence matrices of directed graphs, with certain right hand sides. ~~We~~ also derive, as a special case, the compact linear characterization of the two terminal Steiner tree polytope given in Ball, Liu and Pulleyblank ~~[3]~~.

1. The Result

Consider a collection of nonempty polyhedra in $\mathbb{R}^n$ of the form

$$P_i := \{ x \in \mathbb{R}^n \mid A^i x = d^i,\ x \geq 0 \}, \qquad i \in T,$$

where T is a finite set with $|T| = t \geq 2$, and for $i \in T$, A^i is an $m_i \times n$ matrix and d^i is an m_i-vector.

It is known [1] (see [2] for a published version) that clconv $\cup (P_i : i \in T)$, the (closed) convex hull of the union of the polyhedra P_i, is the set of points $x \in \mathbb{R}^n$ that have an extension $(x, x^1, \ldots, x^t, \lambda) \in \mathbb{R}^{n+tn+t}$ satisfying the conditions

$$(1) \qquad \begin{aligned} x - \Sigma(x^i : i \in T) &= 0 \\ A^i x^i - d^i \lambda_i &= 0, \qquad i \in T \\ \Sigma(\lambda_i : i \in T) &= 1 \\ x^i,\ \lambda_i &\geq 0, \qquad i \in T \end{aligned}$$

This implies that in all basic solutions to the system (1), $\lambda_i = 0$ or 1, $i \in T$.

The question naturally arises, is there a more compact representation of the convex hull in the case when $A^i = A$ for all $i \in T$, i.e. when the polyhedra are of the form

$$P_i := \{ x \in \mathbb{R}^n \mid Ax = d^i,\ x \geq 0 \},\ i \in T.$$

In particular, let Q be the set of those $x \in \mathbb{R}^n$ that have an extension $(x, \lambda) \in \mathbb{R}^{n+t}$ satisfying

$$(2) \qquad \begin{aligned} Ax - \Sigma(d^i \lambda_i : i \in T) &= 0 \\ \Sigma(\lambda_i : i \in T) &= 1 \\ x \geq 0,\ \lambda_i \geq 0,\ i &\in T, \end{aligned}$$

and let $C :=$ cl conv $\cup (P_i : i \in T)$.

It is easy to see that $C \subseteq Q$, since (2) can be obtained from (1) (when $A^i = A$, $i \in T$) by left-multiplying the first equation with A, and then adding

to it all remaining equations except for the last one. However, $C \supseteq Q$ is not true in general.

Jeroslow [5] gave a sufficient condition for $C = Q$ to be true. Blair [4] gave two weaker sufficient conditions, one of which is also necessary when a certain requirement is satisfied. He also showed that recognizing whether $C = Q$ is NP-hard. We give a new sufficient condition for $C = Q$ which is sometimes easier to recognize.

We will denote by $d(\lambda)$ the convex combination of the right hand sides d^i with weights λ_i, $i \in T$; i.e., for $\lambda_i \geq 0$, $i = 1,\ldots,t$, such that $\Sigma(\lambda_i : i \in T) = 1$, $d(\lambda) = \Sigma(d^i\lambda_i : i \in T)$. Also, we will assume that A is of full row rank.

Theorem 1. *$C = Q$ if for every $m \times m$ nonsingular submatrix B of A and every convex combination $d(\lambda)$ of the vectors d^i, $i \in T$,*

(3) $\quad B^{-1}d(\lambda) \geq 0$ *implies* $B^{-1}d^i \geq 0$ *for all* $i \in T$ *such that* $\lambda_i > 0$.

Proof. Let (1') be the system obtained from (1) by replacing each A^i by A. Since $C \subseteq Q$ is always true, we have to show only that $Q \subseteq C$ if (3) holds for every B and λ. Suppose (3) holds and let $\bar{x} \in Q$. W.ℓ.o.g., assume $(\bar{x}, \bar{\lambda})$ is a basic feasible solution to (2). Then there exists an $m \times m$ nonsingular submatrix B of A such that $\bar{x} = (\bar{x}_B, 0)$, where $\bar{x}_B = B^{-1}d(\bar{\lambda}) \geq 0$. From (3), $\bar{\lambda}_i > 0$ implies $B^{-1}d^i \geq 0$. Letting $\bar{x}^i = (\bar{x}^i_B, 0)$, where $\bar{x}^i_B = B^{-1}d^i\bar{\lambda}_i$, we have that $(\bar{x}, \bar{x}^1, \ldots, \bar{x}^t, \bar{\lambda})$ satisfies (1'), hence $\bar{x} \in C$. ||

The assumption that A is of full row rank is not essential. If A does not have this property, we replace it by any of its maximal n-column submatrices of full row rank.

2. An Application: Unions of (Multiple) Network Polyhedra

Next we discuss an important class of problems for which condition (3) of Theorem 1 is always satisfied.

We will say that P_i, $i \in T$, are <u>network polyhedra</u>, if A is the node-arc

incidence matrix of a directed graph, and each d^i has one component equal to some positive number v, one component equal to −v, and all other components equal to 0. If v = 1, the network polyhedra are <u>path polyhedra</u>. The P_i, $i \in T$, will be called <u>multiple network polyhedra</u> (<u>multiple path polyhedra</u>) if A is a block-diagonal matrix with each diagonal block the node-arc incidence matrix of a directed graph, and each d^i has a subvector for each block of A, with exactly one positive and one negative component equal in absolute value (equal to 1 in absolute value), all remaining components being 0. In other words, A and d^i, $i \in T$, are of the form

$$A = \begin{pmatrix} A^1 & & \\ & \ddots & \\ & & A^q \end{pmatrix}, \qquad d^i = \begin{pmatrix} d^{i1} \\ \vdots \\ d^{iq} \end{pmatrix}, \quad i \in T.$$

where the blanks are zero matrices.

Theorem 2. *$C = Q$ if the polyhedra P_i, $i \in T$, are (multiple) network polyhedra, and for each $k \in \{1,\ldots,q\}$, either the negative or the positive component of d^{ik} is in the same position for all $i \in T$.*

The proof of Theorem 2 will use the following auxiliary result. Let A be the node-arc incidence matrix of a directed graph G = (V,E) and for $v \in V$, let A_v be the matrix obtained from A by deleting row v. It is well known that if G is (weakly) connected, the rank of A and of A_v is $|V| - 1$, and that any $(|V|-1) \times (|V|-1)$ nonsingular submatrix of A_v can be made lower triangular by appropriate row and column permutations. In addition, every such matrix has the following property.

Lemma 3. *Let $B = (b_{ij})$ be a $(|V|-1) \times (|V|-1)$ lower triangular submatrix of A_v. Then $B^{-1} = (\beta_{ij})$ is lower triangular and for $i = 1,\ldots,|V|-1$, $j < i$, every nonzero β_{ij} has the same sign as b_{ii}.*

Proof. Let the rows and columns of B be indexed by $M = \{1,\ldots,m = |V|-1\}$, and for R, $C \subseteq M$, let B_R^C be the submatrix of B with rows and columns in R and C, respectively.

We use induction on m. The statement is trivially true for $m = 1$. Suppose it is true for $m = 1,\ldots,k$ and let $m = k + 1 \geq 2$. Then, denoting $K = \{1,\ldots,k = m-1\}$,

$$B^{-1} = \left(\begin{array}{c|c} B_K^K & 0 \\ \hline B_m^K & b_{mm} \end{array}\right)^{-1} = \left(\begin{array}{c|c} (B_K^K)^{-1} & 0 \\ \hline -b_{mm}^{-1}\ B_m^K(B_K^K)^{-1} & b_{mm}^{-1} \end{array}\right).$$

By the induction hypothesis, for $i = 1,\ldots,k$ every nonzero entry in row i of $(B_K^K)^{-1}$ has the same sign as b_{ii}. Hence the claim of the Lemma is true for the first $m-1$ rows of B^{-1}. On the other hand, the i-th entry of the row vector B_m^K, if nonzero, has the sign opposite to that of b_{ii}(from the node-arc incidence relation). Thus $-B_m^K(B_K^K)^{-1} \geq 0$ and therefore every nonzero entry in row m of B^{-1} has the same sign as b_{mm}^{-1}, hence b_{mm}. This completes the induction.||

Proof of Theorem 2. Since each matrix A^k, $k = 1,\ldots,q$, has a redundant row, we delete from each A^k the row corresponding to the unique nonzero entry of d^{ik} which is in the same position for each $i \in T$. We then delete the corresponding entry from each d^{ik}, $i \in T$; and if for some $k \in \{1,\ldots,q\}$ the remaining nonzero entry of each d^{ik}, $i \in T$, is negative, we also change the sign of A^k and of each d^{ik}, $i \in T$. Let $\tilde{A}^k$ and $\tilde{d}^{ik}$ be the resulting matrix and vectors. When we have applied this operation to each block of A, the resulting vectors $\tilde{d}^i$, $i \in T$, are all nonnegative and the resulting matrix $\tilde{A}$ is of full row rank. The polyhedra P_i are now defined by the systems $\tilde{A}x = \tilde{d}^i$, $x \geq 0$, $i \in T$.

Let m be the number of rows of $\tilde{A}$. Then any $m \times m$ nonsingular submatrix B of $\tilde{A}$ can be brought by row and column permutations to the form

$$B = \begin{pmatrix} B_1 & & \\ & \ddots & \\ & & B_q \end{pmatrix},$$

where each B_k in a square lower triangular matrix and the blanks are zeros.

Let B be any such matrix; then its inverse is

$$B^{-1} = \begin{pmatrix} B_1^{-1} & & \\ & \ddots & \\ & & B_q^{-1} \end{pmatrix}.$$

From Lemma 3, B^{-1} is lower triangular and in every row $i \in \{1,\ldots,m\}$ of B^{-1}, all nonzero entries have the sign of b_{ii}. Since $d^i \geq 0$, $i \in T$, for any $\lambda \in R^{|T|}$, $\lambda \geq 0$, such that $\Sigma(\lambda_i : i \in T) = 1$, $B^{-1}d(\lambda) \geq 0$ implies that every diagonal element b_{ii} such that $\lambda_i > 0$ must be 1 (i.e. positive). But then $B^{-1}d^i \geq 0$ for all $i \in T$ such that $\lambda_i > 0$, i.e. the polyhedra P_i satisfy condition (3) of Theorem 1; therefore $C = Q$.||

3. Another Application: Two Terminal Steiner Tree Polyhedra

As a further application, we dicuss the case of two terminal Steiner tree polyhedra and derive from our result the compact linear characterization obtained (by other means) by Ball, Liu and Pulleyblank [3]. Let $G = (V,E)$ be an arc-weighted *directed graph with three distinguished nodes:* a source (root) s, and two sinks (terminals), t and r. A two terminal Steiner tree in G is a minimum-weight arborescence rooted at s and containing nodes t and r. Such an arborescence is the union of three directed paths, from s to some node v, from v to t and from v to r, where v may or may not be distinct from s, t and r. However, the converse is not true, i.e. the union of directed paths from s to v, from v to t and from v to r, is not necessarily a rooted arborescence: it *is* one if and only if the three paths are node-disjoint. When this is not the case, the union of the three paths is called a Steiner net.

For any $v \in V$, the incidence vectors of directed paths from s to v in G are the extreme points of the polyhedra defined by the system $Ax = e_s - e_v$, $x \geq 0$ where A is the node-arc incidence matrix of G and for $i \in V$, e_i is the i-th unit vector in $\mathbb{R}^{|V|}$. Also, the incidence vectors of directed paths from

v to t and from v to r in G are the extreme points of the polyhedra defined by the systems $Ax = e_v - e_t$, $x \geq 0$ and $Ax = e_v - e_r$, $x \geq 0$, respectively.

Now let P(v) be the set of those $x \in \mathbb{R}^{|E|}$ for which there exist vectors $x^s, x^t, x^r \in \mathbb{R}^{|E|}$ satisfying

$$
(4) \qquad
\begin{array}{rcl}
Ax^s \qquad\qquad & = & e_s - e_v \\
\qquad Ax^t \qquad & = & e_v - e_t \\
\qquad\qquad Ax^r & = & e_v - e_r \\
-Ix^s - Ix^t - Ix^r + Ix & = & 0 \\
x^s,\ x^t,\ x^r \geq 0. & &
\end{array}
$$

It is not hard to see that the extreme points of P(v) are precisely those $x \in \mathbb{R}^{|E|}$ such that $x = x^s + x^t + x^r$ for some incidence vectors x^s, x^t and x^r of directed paths in G from s to v, from v to t and from v to r, respectively. Indeed, since x is unconstrained in sign, the last matrix equation of (4) does not impose any constraint on x^s, x^t, and x^r. Therefore, x is an extreme point of P(v) if and only if each of x^s, x^t, and x^r is an extreme point of the polyhedron defined by $x^s \geq 0$ and the first, $x^t \geq 0$ and the second, $x^r \geq 0$ and the third matrix equation of (4), respectively.

To get from (4) a system of full row rank, we delete row s from the first matrix equation, row t from the second, and row r from the third one, and we also change the signs in the first matrix equation. We thus obtain the system

$$
(5) \qquad
\begin{array}{rcl}
-A_s x^s \qquad\qquad & = & e_v \\
\qquad A_t x^t \qquad & = & e_v \\
\qquad\qquad A_r x^r & = & e_v \\
-Ix^s - Ix^t - Ix^r + Ix & = & 0 \\
x^s,\quad x^t,\quad x^r \geq 0, & &
\end{array}
$$

equivalent to (4); and we can write for $v \in V$,

$$
P(v) = \left\{ x \in \mathbb{R}^{|E|} \;\middle|\; \begin{array}{l} \exists\, x^s, x^t, x^r \in \mathbb{R}^{|E|} \text{ such that} \\ (x^s, x^t, x^r, x) \text{ satisfies (5)} \end{array} \right\}
$$

Consider now the two-terminal Steiner tree problem in G. If the weights c_{ij} are nonnegative, then this problem is

$$\min \{ cx \mid x \in \operatorname{conv} \cup (P(v) : v \in V) \},$$

since the minimum is always attained for some x that is the incidence vector of an arborescence rooted at s and continuing nodes t and r. (If the costs are arbitrary and there exists a negative-cost directed cycle, the above problem has no finite minimum. If G has no negative-cost cycles but some of the costs are negative, the optimum may occur for a Steiner net instead of a Steiner tree, and some arcs may have a flow greater than one).

Theorem 4. *The (closed) convex hull of the union of $P(v)$ for all $v \in V$ is the set of those $x \in \mathbb{R}^{|E|}$ for which there exist vectors x^s, x^t, $x^r \in \mathbb{R}^{|E|}$ and $\lambda \in \mathbb{R}^{|V|}$ satisfying the system*

$$\begin{array}{lrcl} & -A_s x^s \qquad\qquad\qquad & -I\lambda & = 0 \\ (6) & A_t x^t \qquad\quad & -I\lambda & = 0 \\ & A_r x^r & -I\lambda & = 0 \\ & -Ix^s - Ix^t - Ix^r + Ix & & = 0 \\ & & 1\lambda & = 1 \\ & x^s,\ x^t,\ x^r \geq 0, & \lambda & \geq 0. \end{array}$$

Proof. The system (6) can be rewritten as

$$\begin{array}{ll} & My - L\lambda = 0 \\ (7) & 1\lambda = 1 \\ & y,\ \lambda = 0 \end{array}$$

where $L = (I, I, I)^T$ and $y = (x^s, x^t, x^r, x) \in \mathbb{R}^{4|E|}$. We will show that condition (3) of Theorem 1 is satisfied for (7) viewed as an instance of (2). The matrix M has $3|V| - 3 + |E| := p$ rows and is of full row rank. Any $p \times p$ nonsingular submatrix of M can be brought by row and column permutations to

the form

$$B = \begin{pmatrix} B_s & & & \\ & B_t & & \\ & & B_r & \\ J_s & J_t & J_r & I \end{pmatrix},$$

where B_s, B_t and B_r are $(|V| - 1) \times (|V| - 1)$ lower triangular submatrices of A_s, A_t and A_r, respectively, J_s, J_t and J_r are $|E| \times (|V| - 1)$ submatrices of I, I itself is the identity matrix of order $|E|$, and the blanks are zero matrices. Clearly, B^{-1} is of the form

$$B^{-1} = \begin{pmatrix} B_s^{-1} & & & \\ & B_t^{-1} & & \\ & & B_r^{-1} & \\ J_s^* & J_t^* & J_r^* & I \end{pmatrix}$$

for some unspecified J_s^*, J_t^* and J_r^*.

From Lemma 3, each diagonal block of B^{-1} is lower triangular and every nonzero entry in any of the first $3|V|-3$ rows of B^{-1} has the same sign as the corresponding diagonal element of B. It then follows that for $B^{-1}L\lambda \geq 0$ to be satisfied for any $\lambda \geq 0$ (with $1\lambda = 1$), it is necessary that every diagonal entry of B be +1. But then $B^{-1}L\lambda \geq 0$ implies $B^{-1}\ell^i \geq 0$ for every column ℓ^i of L such that $\lambda_i > 0$, i.e. condition (3) of Theorem 1 is satisfied.||

Theorem 4 asserts that all basic solutions to the system (6) are integer, and thus (6) represents a compact (polynomial-sized) linear characterization of the two terminal Steiner tree polytope. This characterization is not new; in fact, it is due to Ball, Liu and Pulleyblank[3]. However, our proof of its validity is new, and puts this problem into the more general context of unions of polyhedra whose convex hull has the simple representation (2).

It should be mentioned that the approach discussed here can be extended to k-terminal Steiner tree polytopes; the resulting formulation involves the union of a number of polytopes exponential in k, but polynomial in $|V|$ for fixed k.

Acknowledgement

Thanks are due to Charles Blair for pointing out an error in an earlier draft of this paper.

References

[1] E. Balas, "Disjunctive Programming: Properties of the Convex hull of Feasible Points." MSRR No. 348, July 1974.

[2] E. Balas, "Disjunctive Programming and a Hierarchy of Relaxations for Discrete Optimization Problems." *SIAM Journal on Algebraic and Discrete Methods, 6,* 1985, 466-486.

[3] M. O. Ball, W. Liu and W. R. Pulleyblank, "Two Terminal Steiner Tree Polyhedra." Report No. 87466-OR, Institut fur Okonometrie und Operations Research, University of Bonn, April 1987.

[4] C. Blair, "Representation for Multiple Right Hand Sides." Business Administration Department, University of Illinois at Champaign, 1987.

[5] R. G. Jeroslow, "A Simplification for Some Disjunctive Formulations." Georgia Institute of Technology, December 1985.